Results demonstrate the importance of sensitivity analyses and thoughtful consideration of the metric and benchmark selection.

PRM45 HEALTH RESOURCE UTILIZATION OF LUPUS NEPHRITIS PATIENTS – A COMPARISON OF RESULT ACROSS CASE IDENTIFICATION ALGORITHMS

Yung-Lin Clark DWJ1, SCG2

Biogen IDEC, Westwood, MA, USA

OBJECTIVES: Lupus nephritis (LN) is a severe complication of systemic lupus erythematosus (SLE). While LN has no designated ICD-9 code, various approaches have been used to identify patients with LN in administrative claims databases. The objective was to compare health resource utilization of LN patients using different algorithms in a single data source. METHODS: This study used the Impact Data, a large commercial insurance claims database. SLE patients were identified using ICD-9 code 710.0 from ≥2 outpatient or ≥1 inpatient claims from 01/2004 to 06/2011. SLE patients with LN were further identified under four different algorithms: (A) ≥1 renal diagnosis, (B) ≥2 renal diagnoses, (C) ≥3 renal diagnoses, and (D) ≥3 renal diagnoses plus ≥3 nephrologist visits. Health resource utilization and expenditure of outpatient/emergency department (ED) visits, hospitalizations, and prescriptions were examined for 12 months post index date of first renal diagnosis. RESULTS: A total of 93,957 patients were diagnosed with SLE. Among them, 24,357, 11,054, 8,895, and 6,307 cases had LN using algorithms A-D. LN cases identified by algorithms A-D had similar mean age (48.3, 46.7, 46.3, and 45.7 years) and gender distribution (83.1, 83.1, 82.7, and 81.8% females). LN patients from different algorithms also had similar annual frequency of outpatient visit (35.9, 41.0, 41.1, and 42.1), ED visit (1.4, 1.4, 1.4, and 1.5), and prescription (6.8, 7.3, 7.2, and 7.4). The annual medical expenditures were $33,176, $36,974, $36,241, and $38,883 for algorithms A, B, C, and D, respectively. CONCLUSIONS: Our results support that when studying health resource utilization, the results do not differ significantly based on the number of renal diagnoses codes. There is a difference in outcomes when using different algorithms. However, the number of patients was high; however, in the case where specialty information is either unavailable or unreliable, using algorithms A-C proved equally reliable in an administrative claims database.

PRM46 ANALYSIS OF HEALTH CARE COSTS CONTAINING A LARGE PROPORTION OF $0 DATA USING TRADITIONAL AND ZERO-INFLATED GAMMA REGRESSION MODELS

Kern DM1, Wasser T1

HealthCore, Inc., Wilmington, DE, USA

OBJECTIVES: To compare traditional and zero-inflated gamma regression models for the analysis of health care costs in an administrative claims database. GAMMA regression models are widely used for the analysis of health care cost data and are appropriate for analyzing mean costs within patients incurring ≥$0. However, little research is available for the appropriateness of gamma regression models on a data with a high proportion of $0 costs. METHODS: This study used the HealthCore Integrated Research Environment to analyze hospitalization costs within 3,049 bipolar patients initiating antipsychotic medications (Drug X=2,398 patients, Drug Y=651 patients). Results of a traditional gamma model and a zero-inflated gamma model were compared. The zero-inflated model used SAS procedure NL MIXED to perform logistic regression modeling the probability of having a hospitalization. GAMMA regression modeling mean costs within patients having a hospitalization, and combining the two models to analyze differences in overall mean cost. RESULTS: There were 18.3% Drug X patients with ≥1 inpatient hospitalization, compared with 12.1% of Drug Y. Mean costs were higher for Drug X within only patients having an event ($17,721 vs. $11,425) and including patients with no event ($3,237 vs. $1,387). GAMMA regression found no differences in the overall mean cost to be significant (p=0.048). CONCLUSIONS: Differences in mean costs between traditional and zero-inflated GAMMA models showed similar results (mean-diff=$1,850, CI=[$1,237,$2,464], p<0.0001). Consistent results between the models were also seen when adjusting for patient demographics, comorbidities, and prior medications: traditional gamma model (mean-diff=$2,173, CI=[$716,$2,297], p<0.0001), zero-inflated gamma model (mean-diff=$2,448, CI=[$597,$7,160], p<0.0001). CONCLUSIONS: When analyzing health care cost data containing ≥80% $0 costs, little difference was seen between traditional gamma regression and zero-inflated gamma regression models. The zero-inflated model uses complex coding, and requires advanced knowledge of statistical methods and SAS programming. Within data containing a large proportion of zero costs, traditional gamma regression is appropriate for analyzing differences in mean costs.

PRM48 DIAGNOSTICS FOR CHECKING THE GAMMA DISTRIBUTION ASSUMPTION IN GENERALIZED LINEAR MODELS USED FOR MODELING HEALTH CARE COSTS

Jain S1, Pan KY2, Durkendahl S3

1Truven Health Analytics, Bethesda, MD, USA, 2Truven Health Analytics, Cambridge, MA, USA, 3Truven Health Analytics, Washington, DC, USA

OBJECTIVES: 1) To study the relative performance of various graphical and analytical methods to assess the assumption of an underlying gamma distribution for a commonly used model in the examination of health expenditures (GLMAGALL – a generalized linear model assuming a gamma distribution with a log link relating the mean costs to a corresponding set of predictors). The objective was to investigate each diagnostic’s ability to discriminate between various assumed distributions (e.g., gamma versus log-normal) that may be potentially used in health care cost analysis. METHODS: Data sets will be created via Monte Carlo simulation of gamma (varying the shape and scale parameters) and other distributions, as well as, by altering the values for a small set of predictors. The results of the diagnostic study will be illustrated in a graphical and/or tabular format. RESULTS: Graphical diagnostics afford the analyst the ability to see subtle or dramatic departures from the model’s distributional assumptions that might not be as obvious by using an analytical model that provides a single summary statistic. CONCLUSIONS: The performance of a diagnostic procedure to assess the presence of a gamma distribution is a cost of discriminate model or its ability to be used in the case where one distribution and another is important; however, other factors must be considered before an analyst makes his or her final choice. The ease of executing the technique, its relative clarity of interpretation, and availability in a software package (without having to perform extensive programming beyond what is provided by a standard statistical package) must all be considered to ensure that model adequacy testing may be performed readily so that the choice of a distribution for an expenditure model may be considered sound.

RESEARCH ON METHODS – Databases & Management Methods

PRM50 USING TEXT MINING OF ELECTRONIC MEDICAL RECORDS TO IDENTIFY KRAS TESTING STATUS IN MCRC PATIENTS

Miller FY1, Walker MS2, Landman Blumberg P1, Cuyun Carter G1

1ACORN Research, LLC, Memphis, TN, USA, 2Truven Health Analytics, Bethesda, MD, USA

OBJECTIVES: To develop algorithms identifying if metastatic colorectal cancer (mCRC) patients were tested for KRAS (a tumor biomarker of EGFR-inhibitor response) using text documents (e.g., physician progress notes) within electronic medical records (EMR). METHODS: The sample consisted of 1,385 mCRC patients from the ACORN Data Warehouse. 300 patients were randomly selected for chart reviews (R blinded to patient grouping (N=150), or testing (N=50), or testing datasets (N=100); 1,085 patients comprised a scoring dataset. Counts of terms in text-based content of patient EMRs were used to develop models predicting KRAS testing status. RESULTS: Traditional models were unable to predict KRAS testing in the training sample. Decision tree (DT), random forest (RF), and adaptive boosting (AB) models performed best when applied to validation data not used in the earlier model development process. RF outperformed DT and AB. RF was the only model to produce a kappa > 0.80 (within rounding) for both the validation and testing datasets. It also produced the highest kappa in the testing dataset (kappa=0.79), as well as fewer false negatives. RF was used to score the remaining 5,015 patients. All patients predicted “tested” and a random sample of patients predicted “not tested” underwent chart review. The model correctly predicted KRAS “tested” 482/500 times (PPV=96.4%) and “not tested” 196/200 times (NPV=98.0%). CONCLUSIONS: Text mining yielded highly accurate classification of KRAS testing status among mCRC patients. Review of the small number of misclassified cases of KRAS testing identified ways to improve the model’s accuracy. These results may inform future research and reduce the need for labor-intensive and costly full chart review by human coders.

PRM51 THERE BUT FOR GRACE? A VALIDATED SCREENING TOOL FOR QUALITY OBSERVATIONAL STUDIES OF COMPARATIVE EFFECTIVENESS

Drewer EA1, Velangas P1, Westrich KB2, Dubois RW1

1Quintet Outcome, Cambridge, MA, USA, 2National Pharmaceutical Council, Washington, DC, USA

OBJECTIVES: To be able to identify observational studies of good enough quality for decision support by validating a set of screening questions to qualify study samples that are likely to produce reasonably accurate and unbiased estimates of comparative effectiveness (CE). METHODS: An 11-item checklist was developed through literature review and consultation with experts from ISPOR, ISPE, payer groups, industry, and academic. Item content covers four quality domains: comparability of subjects, information about the exposure or intervention, outcome measurement, and statistical analysis, which are metrics similar to those used in assessing observational study quality for systematic reviews. Checklist items were tested using studies of drugs, medical devices and medical procedures. We focused on research quality, not applicability to any decision. A fundamental challenge was to find a gold standard against which to test checklist items. 113 volunteers from 5 continents each rated >3 articles (N=280 assessments) from three validation sets of studies that 1) had quality assessments published in systematic reviews; 2) were assessed for quality by one of nine advisors from academic and payer groups; or 3) were assessed for quality by two of the nine advisors. RESULTS: Expert reviews uncovered an unsettling lack of agreement about what “good” looks like, especially in situations that lack context, with 52% concordance (5 experts, 23 assessments). The single best performing checklist item, data quality vs. the validity output, or concordance 0.67 for positive predictive value in 4 of 6 samples and > 0.67 for negative predictive values in all 6 samples. Another high scoring question, sensitivity analysis, was a positive predictive value 0.63 and has no samples. CONCLUSIONS: This quantitative study shows that many content items recommended by experts do not consistently distinguish high quality observational CE studies.

PRM52 ASSOCIATION BETWEEN CARDIOVASCULAR BIOMARKERS LEVELS AND CIGARETTE SMOKING AMONG CURRENT SMOKERS, FAST SMOKERS AND NON SMOKERS USING NHANES 2007-2010

Saxena K1, Liang Q2, Muhammad-Kah R3, Sarkar M2

1Virginia Commonwealth University, Richmond, VA, USA, 2Altria Client Services Inc., Richmond, VA, USA
OBJECTIVES: To investigate the usefulness of National Health and Nutrition Examination Survey (NHANES) data for exposure estimates. To illustrate its utility, the exposure was serum biomarkers of cardiovascular disease (WC, Apo lipoprotein, C-reactive protein, HDL, LDL, total cholesterol) among current, past, and non smokers was examined using NHANES 2007-2010. METHODS: Data were obtained from NHANES 2007 to 2010. The study sample of 11,960 respondents, was divided into prevalent cases where cotinine on smoking and had complete laboratory values for their biomarkers measurement. The population was categorized as current, past, and non smokers. The exposure variable was serum cotinine concentration and the outcome variables were serum levels of the biomarkers listed above. Weighted survey linear regression was used to estimate the association between cotinine concentration and biomarker levels. We also tested the models for different levels of covariates, age, race, and sex. RESULTS: RESULTS: The levels of WC were F-value: 38.78, P-value: <0.0001 and HDL (F-value: 26.4; P-value: <0.0001) showed higher association with cotinine levels than rest of the biomarkers. The R² for the models ranged from 0.039-0.261. Higher levels of WBC and lower levels of HDL were observed among current smokers relative to past and non smokers. BMI showed a high association with most of the biomarkers. The odds of lower LDL were significantly higher for 21-35 years age group relative to the >65 years age group.

PRM53 CLAIMS DATA ALGORITHMS FOR IDENTIFYING INCIDENT COLORECTAL CANCER (CRC) CASES AND CANCER DISEASE STAGE: A CRITICAL REVIEW OF THE LITERATURE

Foley K1, Shi N1, Girvan A2, Ward KC1, Lipscomb J1
1Truven Health Analytics, Cambridge, MA, USA, 2Eli Lilly and Company, Indianaplis, IN, USA, 3Emory University, Rollins School of Public Health, Atlanta, GA, USA, 4Emory University, Atlanta, GA, USA

OBJECTIVES: Administrative claims data are commonly used to study CRC treatment patterns and outcomes. We critically review existing algorithms for identifying incident CRC cases and disease stage at diagnosis within claims data. METHODS: A literature search (1989-2012) using the terms: cancer [ti] AND (administrative OR claims) AND (assess [ti] or  assessment [ti] or  predict [ti] or  diagnostic code [ti]) identified 76 articles with 27 testing algorithms. Of the 27, three provided algorithms for identifying incident BC cases and four algorithms classified BC disease stage at diagnosis. We examine the positive predictive value (PPV) of each algorithm and suggest revisions for improving the PPV. RESULTS: Setoguchi evaluated four algorithms for identifying CRC patients. The first and most restrictive required combinations of ICD-9 and treatment codes. The second used two diagnoses within two months, the third combined the first and second, and the fourth required one diagnosis. Ramsey used one or more ICD-9 codes, the same as Setoguchi algorithm 4. The PPV ranged from 0.039-0.261. Higher levels of WBC and lower levels of HDL were observed among current smokers relative to past and non smokers. BMI showed a high association with most of the biomarkers. The odds of lower LDL were significantly higher for 21-35 years age group relative to the >65 years age group. Non-Hispanic blacks had a significantly lower LDL than non-Hispanic whites. Females had significantly higher LDL than males. These results were consistent with that reported in the literature. CONCLUSIONS: A statistically significant association was observed between the biomarker exposure (serum cotinine) and WBC and HDL cholesterol. There were also significant differences in the association within the different covariate levels.

PRM54 INFlixIMAB COMPLIANCE ESTIMATES FROM MEDICAL CHART AND ADMINISTRATIVE CLAIMS DATABASES: A METHODOLOGICAL COMPARISON

Kizziar C1, Traczyk J1, Kandel P1
1Health Analytics, Columbia, MD, USA, 2Health Analytics, LLC, Columbia, MD, USA

OBJECTIVES: The purpose of this study was to compare the medication compliance estimates among a sample of infliximab-treated Crohn’s disease (CD) patients using medical chart and administrative claims (a claims) data. METHODS: A Mid-Atlantic regional health plan provided claims data for their CD population during calendar years 2006-2010. Claims were aggregated by a date, dosage, and administered quantity were all available. Approximately 16,607 hypertensive patients had a stable dose of 5 mg/kg over the course of their first episode of care, with a mean maintenance interval of 60 days. Dosing information from patient charts will be used to estimate dosage from claims where charts are not in present to help create an algorithm for estimating dose from administrative claims only. CONCLUSIONS: Claims data were less equipped to estimate biologic treatment compliance, while the chart data provides more independent sources to back biologic dosing patterns. However, information from claims augmented charts, and provided details on a much larger population of biologic patients, highlighting the importance of utilizing both the get a better depiction of treatment, compliance, and costs.

PRM55 CLAIMS DATA ALGORITHMS FOR IDENTIFYING INCIDENT BREAST CANCER (BC) CASES AND CANCER DISEASE STAGE: A CRITICAL REVIEW OF THE LITERATURE

Foley K1, Shi N1, Girvan A1, Ward KC1, Lipscomb J1
1Truven Health Analytics, Cambridge, MA, USA, 2Eli Lilly and Company, Indianapolis, IN, USA, 3Emory University, Rollins School of Public Health, Atlanta, GA, USA, 4Emory University, Atlanta, GA, USA

OBJECTIVES: Administrative claims data are used to study BC treatment patterns and outcomes. We critically review existing algorithms for identifying incident BC cases and disease stage at diagnosis within claims data. METHODS: A literature search (1989-2012) using the terms: cancer [ti] AND (administrative OR claims) AND (assess [ti] or  assessment [ti] or  predict [ti] or  identify [ti] or  identification [ti] or validate [ti] or validation [ti]) identified 76 articles with 27 testing algorithms. Of the 27, three provided algorithms for identifying incident BC cases and four algorithms classified BC disease stage at diagnosis. We examine the positive predictive value (PPV) of each algorithm and suggest revisions for improving the PPV. RESULTS: To identify incident BC, Warren used ICD-9 and treatment codes across inpatient and outpatient settings (PPV=80%). Freeman used a logistic regression model with 36 diagnostic and procedural indicators and Naginther used a 4-step process using ICD-9 and treatment codes (PPV=93%). For disease stage, the Yuen and Cooper papers used diagnostic codes to distinguish regional from distant spread of cancer, with PPVs below 60%. Smith used diagnostic procedures, and employed equations to create equations for distinguishing stage IV from all other patients and stage III from patients with stage I/II. Trade-offs between sensitivity and PPV are made using cut-points from BC models to classify patients into disease stage. Nordstrom’s algorithm identified metastatic disease using three components: a diagnosis code for secondary neoplasm OR any metastatic chemotherapy agent OR no G-code for non-metastatic disease AND G-code for metastatic disease (PPV=81%). CONCLUSIONS: For identifying BC, we recommend updating the Naginther algorithm with newer codes and fewer criteria for older patients likely to receive less aggressive care. For disease stage, we recommend updating the algorithm with additional codes from Nordstrom and testing higher cut-points to maximize PPV.

PRM56 OSTEORHISTIOMATOSIS IS AN INDEPENDENT RISK FACTOR FOR HIGHER EVENT RATE OF MAJOR ADVERSE CARDIOVASCULAR OUTCOMES IN HYPERTENSIVE PATIENTS – FROM 15-YEAR TAIWAN NHIRD DATABASE

Chang Bing Shou, Chuan Memorial Hospital, Changhua City, Taiwan

OBJECTIVES: Osteorhismatosis andatherosclerosis are both chronic inflammatory diseases.However, the association between osteorhismatosis and the rate of major adverse cardiovascular events (MACE) has not been sought (literature). METHODS: Using the retrospective, observational, longitudinal study design, we evaluate the association between osteorhismatosis and the rate of MACEs in the Taiwan essential hypertensive cohort. RESULTS: From National Health Insurance Research Dataset (NHIRD). Patients with HTN and free of pre-defined MACEs (MI, stroke, CHF, ESRD and PVD) in the entire course were enrolled as control group. In contrast, those both with HTN and OA and free of previous MACEs were enrolled as study group. Totally up to 56,607 hypertensive patients (aged 30-60 years) without previous MACEs at the first year (in 1996) were identified. There were 23,530 (41.6 %) patients with concomitant diagnosis of OA with those without OA (MI: 2.37% vs. 1.64 %, p <0.001; stroke: 3.30% vs.3.33%, p=0.0004; CHF: 1.60% vs.1.37%, p=0.025; PVD: 10.24% vs. 4.36%, p<0.0001, respectively), except for ESRD (1.05% vs. 1.33%, p=0.0032). After adjusted for birth year and sex, the adjusted ORs (95% confidence interval (CI) and p-values) for MACE at the crude MACE rates were significantly higher in hypertensive patients with OA than those without OA (MI: 2.37% vs. 1.64 %, p <0.001; stroke: 3.30% vs.3.33%, p=0.0004; CHF: 1.60% vs.1.37%, p=0.025; PVD: 10.24% vs. 4.36%, p<0.0001, respectively). For patients with OA and HTN were all significantly higher by the relative risk of 3.09 (2.69-3.54, p<0.0001) in MI; 2.47 (2.22-2.75, p<0.0001) in stroke; 4.60 (2.92-7.01, p<0.0001) in CHF; 1.75 (1.48-2.08, p<0.0001) in ESRD and 4.77 (4.38- 4.19, p<0.0001) in PVD, respectively. CONCLUSIONS: Results from this study highlight the MACE (MI, stroke, CHF, ESRD and PVD) rates were significantly higher in hypertensive patients with OA than those without it.

PRM57 COMPARISON OF COMMERCIAL INSURANCE DATABASES TO CENSUS DATA FOR AGE, GENDER AND GEOGRAPHIC REGION IN THE UNITED STATES

Wasser T1, Wu Y1, Ycas J2, Tunceli O1, Cziraky MJ1
1HealthCore, Inc., Wilmington, DE, USA, 2AstraZeneca Pharmaceuticals LP, Wilmington, DE, USA

OBJECTIVES: Commercial insurance data, such as MarketScan, are considered non-representative of the United States (US) population because they reflect only working age individuals, and their dependents who are currently employed. If this employment characteristic exists, it should be visible when large commercial administrative databases are compared against US Census demographic data. METHODS: This study compared the HealthCore Integrated Research Database (HIRD) and the MarketScan Database against US Demographic data for geographic region, age and gender. Age groups and geographic regions were ceded to be consistent with